STIC-ILL

@H345.T7

MPL

From: Steadman, David (AU1652)
Sent: Wednesday, December 05, 2001 9:47 AM
To: STIC-ILL
Subject: Literature/ILL Request

Art Unit: 1652
Office: 10D-04
Mailbox: 10C-01 M3
Case Serial #:09/459,573

Please provide the following references:

1) A new family of amino-acid-efflux proteins
Vladimir V. Aleshin, Natalia P. Zakataeva and Vitaliy A. Livshits
Trends in Biochemical Sciences, 1999, 24:4:133-135

2) Zakataeva NP, Aleshin VV, Tokmakova IL, Troshin PV, Livshits VA.
The novel transmembrane Escherichia coli proteins involved in the amino acid efflux.
FEBS Lett. 1999 Jun 11;452(3):228-32.

Thank you,
David Steadman

number: 2635222)], several proteins from *Clostridium perfringens* (including a hyaluronidase), and a putative serine/threonine kinase from *Synechocystis* sp. Many of the bacterial proteins identified are from intracellular pathogens that infect eukaryotic cells and probably are involved in cell invasion.

Threading calculations and model building provide convincing evidence that the N-terminus of the P60 invasion protein has an SH3 fold. The UCLA fold-recognition server[4] predicted that P60_LISGR contains a region that has a fold similar to that of the SH3 domain of the FYN proto-oncogene tyrosine kinase [PDB entry: 1shf (Z = 6.70, which is well above the confidence threshold of 5.0 ± 1)]. In addition, eight out of the ten highest-scoring results had folds homologous to SH3 domains; the two highest scoring – both SH3 domains – had Z scores of >5.0. A second fold-recognition server, THREADER2 (Ref. 5), returned as the two highest-scoring results 1shf (the SH3 domain from the FYN proto-oncogene tyrosine kinase; Z = 7.68) and 1shg (the SH3 domain from α-spectrin; Z = 6.81). Both scores are well above the 'very significant' threshold for THREADER2 (Z = 3.5). The next-best result, 1mjc (the major cold-shock protein 7.4 of *Escherichia coli*), which does not contain an SH3 domain, had a substantially lower score (Z = 3.0).

We built a model of the fragment of P60_LISGR based on the chicken SRC tyrosine kinase[6], using the alignment shown in Fig. 1. All residues buried in the chicken SH3-domain structure correspond to hydrophobic residues (or threonine or glycine residues) in P60_LISGR. An asparagine residue that replaces the conserved proline residue present in the eukaryotic SH3 domains (shown in Fig. 1) is exposed and lies at the bottom of the groove in SH3 domains that bind peptides. The GTPase-activating protein GTPA_RAT and other SH3 homologues have a valine residue at this position, which shows that the proline residue is not essential.

Functional significance. Invasion of eukaryotic cells by most pathogenic bacteria is accompanied by tyrosine phosphorylation, and inhibition of tyrosine phosphorylation impairs invasion by *Listeria monocytogenes*[7]. *Listeria* contain several invasion proteins. Different invasion factors – sometimes in concert – facilitate invasion of different cell types. P60 is important for invasion of epithelial cells[8] and also for survival within the host cell[9]. Indeed, the N-termini of members of the P60 family of invasion proteins are highly conserved among different species of *Listeria*, which implies that this region is functionally important.

The P60 protein itself is thought to be a murine hydrolase[10]. It consists of three domains: the conserved N-terminus, which we suggest is an SH3 domain; a central domain that contains Ser/Thr-rich repeats; and a C-terminal domain, which is homologous to a number of α amylases and starch-degrading enzymes. Species of bacteria that contain homologues of the putative SH3 domain from P60_LISGR are pathogens that invade eukaryotic cells. The SH3 domains of these prokaryotes might therefore have two possible functions: (1) promoting survival of a pathogen within the invaded cell by modulating pathways controlled by SH3 domains; or (2) promoting invasion by binding to receptors on eukaryotic cells.

Conclusions. We have suggested, on the basis of sequence similarity, structural compatibility and function, that P60_LISGR contains an SH3 domain. If this is confirmed, the appearance of SH3 domains in *L. grayi* will extend the range of this important family of proteins to prokaryotes (see Box 1).

Box 1. Note added in proof

After this manuscript was submitted, Bilwes et al. reported the structure of SH3-like domains in CheA, a histidine kinase from the bacterium *Thermotoga maritima* that is homologous to proteins from *Escherichia coli*. The two SH3-like regions in CheA are 'circular permuted' which alters the sequential pattern.
Bilwes et al. (1999) *Cell* 96, 131–141

Acknowledgements

We thank C. Chothia, S. Teichmann and R. Williams for suggestions, and the Wellcome Trust and NHMRC (grant 977144) for support.

References

1 Musacchio, A., Wilmanns, M. and Saraste, M. (1994) *Prog. Biophys. Mol. Biol.* 61, 283–297
2 Falzone, C. J. *et al.* (1994) *Biochemistry* 33, 6052–6062
3 Altschul, S. F. *et al.* (1997) *Nucleic Acids Res.* 25, 3389–3402
4 Fischer, D. and Eisenberg, D. (1996) *Protein Sci.* 5, 947–955
5 Jones, D. T., Taylor, W. R. and Thornton, J. M. (1992) *Nature* 358, 86–89
6 Williams, J. C. *et al.* (1997) *J. Mol. Biol.* 274, 757–775
7 Tang, P., Rosenshine, I. and Finlay, B. B. (1994) *Mol. Biol. Cell* 5, 455–464
8 Hess, J. *et al.* (1966) *Zentralbl. Bakteriol.* 284, 263–272
9 Hess, J. *et al.* (1995) *Infect. Immun.* 63, 2047–2053
10 Wuenscher, M. D. *et al.* (1993) *J. Bacteriol.* 175, 3491–3501

JAMES C. WHISSTOCK

Dept of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, Victoria 3168, Australia.

ARTHUR M. LESK

Dept of Haematology, Cambridge Institute for Medical Research, University of Cambridge Clinical School, Hills Rd, Cambridge, UK CB2 2XY.

A new family of amino-acid-efflux proteins

Analyses of bacterial genome sequences reveal many genes that encode putative membrane proteins. Many known membrane proteins are involved in the transport of compounds into the cell[1,2]. The transporters involved in efflux are less well studied, although they play important roles in resistance to toxic substances, in maintenance of an optimum intracellular concentration of metabolites, and in excretion of some regulatory molecules[3–5].

Homoserine, a metabolic precursor of threonine and methionine, is an important regulator in various bacteria. In *Escherichia coli*, homoserine inhibits $NADP^+$-specific glutamate dehydrogenase (E.C. 1.4.1.4), the enzyme that catalyses the primary reaction in ammonium assimilation[6]. Moreover, homoserine lactone, which is generated from homoserine[7], activates the expression of the σ^s subunit of RNA polymerase, the subunit that provides transcriptional specificity for the groups of genes that are switched on during starvation and/or on entering stationary phase[8]. Accordingly, exogenous homoserine lactone and homoserine suppress the growth of *E. coli* in minimal nutritional media, probably by stimulating expression of σ^s.

Amplification of genes that encode components of systems involved in the efflux of antibiotics, organic solvents and metal ions increases the resistance of bacteria to these substances[3,9–11]. We have found that overexpression of an *E. coli* chromosomal DNA fragment from the 86-min region makes cells resistant to homoserine lactone, homoserine and threonine[12]. The minimum fragment length necessary for producing such a phenotype is 0.8 kb and includes the open reading frame (ORF) f138 (GenBank accession number M87049)[13] and 348 bp of DNA

 PII: S0968-0004(99)01367-5

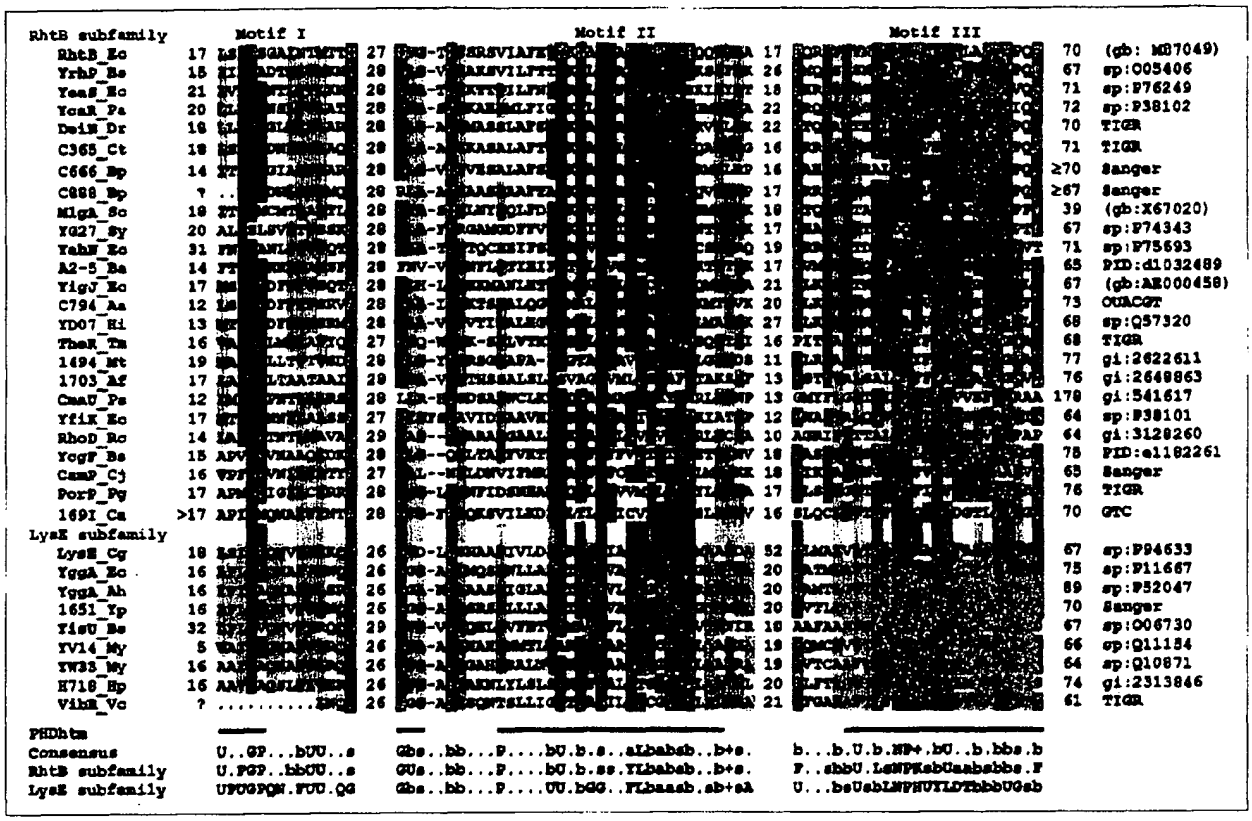


Figure 1

Multiple alignment of RhtB proteins. The fragments listed were selected from >60 sequences on the basis of the maximum dissimilarity in their primary structures. The distances between the motifs and the distances from the protein termini are indicated. Where >50% of sequences have similar or identical residues at a given position, a consensus residue is assigned [a, aromatic residue (F, Y or W); U, bulky aliphatic residue (I, L, V or M); b, bulky aliphatic/aromatic residue (I, L, V, M, F, Y or W); s, small residue (G, S, T or A); +, positively charged residue (K, R or H). Conserved residues are highlighted in colour: red indicates residues that fit the general consensus well; yellow indicates residues that fit the general consensus to a lesser extent; blue indicates residues that fit the RhtB-subfamily consensus; green indicates residues that fit the LysE-subfamily consensus. The positions of predicted transmembrane helices are shown as thick black lines. Accession numbers in databases (gb, GenBank; gi, gene identification; PID, protein identification; sp, SWISS-PROT) or the contributing genome centers for sequences of unfinished genomes (GTC, Genome Therapeutics Corporation; OUACGT, University of Oklahoma Advanced Center for Genome Technology; Sanger, Sanger Centre; TIGR, The Institute for Genomic Research) are indicated in the right-hand column. Feature tables of the items shown in brackets were modified by either shifting the translation-initiation point or partial alteration of the reading frame. Aa, *Actinobacillus actinomycetemcomitans*; Af, *Archaeoglobus fulgidus*; Ah, *Aeromonas hydrophila*; Ba, *Bacillus* sp.; Bp, *Bordetella pertussis*; Bs, *Bacillus subtilis*; Ca, *Clostridium acetobutylicum*; Cg, *Corynebacterium glutamicum*; Cj, *Campylobacter jejuni*; Ct, *Chlorobium tepidum*; Dr, *Deinococcus radiodurans*; Ec, *Escherichia coli*; Hi, *Haemophilus influenzae*; Hp, *Helicobacter pylori*; Mt, *Methanobacterium thermoautotrophicum*; My, *Mycobacterium tuberculosis*; Pa, *Pseudomonas aeruginosa*; Pg, *Porphyromonas gingivalis*; Ps, *Pseudomonas syringae*; Rc, *Rhodobacter capsulatus*; Sc, *Shewanella colwelliana*; Sy, *Synechocystis* sp. PCC 6803; Th, *Thermotoga maritima*; Vc, *Vibrio cholerae*; Yp, *Yersinia pestis*.

upstream of this ORF. Note that a construct that contains only 160 upstream nucleotides does not provide resistance to the above-mentioned amino acids. The upstream sequence does not contain a stop codon in frame with ORF f138. Moreover, one of the ATG codons in this sequence is preceded by a predicted ribosome-binding site. We designated the resultant, extended ORF (62160–61546 bp in M87049) *rhtB*. Disruption of the chromosomal *rhtB* gene causes hypersusceptibility to homoserine lactone and homoserine (V. V. Aleshin, unpublished). The RhtB protein is predicted to be highly hydrophobic and to possess six transmembrane segments.

We have found a set of proteins that are homologous to RhtB in a wide range of prokaryotes that includes proteobacteria, cyanobacteria, bacilli and mycobacteria, and the archaea *Archaeoglobus fulgidus* and *Methanobacterium thermoautotrophicum* (Fig. 1). We performed a PSI-BLAST[14] search of the non-redundant database at the NCBI and gapped BLAST[14] searches of unfinished microbial genomes. A PSI-BLAST search, with an *E*-value threshold of 10^{-3}, retrieved a set of proteins in three iterations – after which the search converged. In a gapped BLAST search, the probabilities of chance matches were estimated for the most-closely related sequences ($p < 10^{-25}$) and the most-distantly related ($p < 10^{-3}$) sequences. Most of the sequences homologous to the RhtB sequence represent hypothetical transmembrane proteins, some of which recently have been included in the UPF0048 family. One, LysE, is the only transporter known to be responsible for the efflux of an amino acid: it conducts lysine in *Corynebacterium glutamicum*[15]. We suggest that RhtB is involved in the efflux of homoserine and threonine in *E. coli*.

We generated unrooted dendrograms by neighbour-joining and maximum-parsimony methods, using the PHYLIP 3.572 package with bootstrap analysis[16]. Dendrograms (not shown) indicate that two different subfamilies exist: an RhtB-related subfamily and a LysE-related subfamily (Fig. 1). Some

genomes encode several paralogs from the two subfamilies (e.g. *Bacillus subtilis*, *E. coli* and *Pseudomonas aeruginosa* encode three, six and 12 paralogs, respectively). Thus, the divergence between the subfamilies is associated with gene duplication rather than with taxonomic diversification and occurred before the divergence of Gram-positive and Gram-negative bacteria.

Multiple alignment by using the MACAW program[17] revealed that three motifs are significantly conserved ($p < 10^{-18}$) in all these proteins: (1) a three-residue motif near the N-terminus (PGP in the RhtB subfamily, and PXGP in the LysE subfamily); (2) an aromatic motif that lies ~60 residues from the N-terminus; and (3) an $FX_7LXNP^K/_HX_2LX_8F$ motif that lies 16–58 residues C-terminal to the second motif (Fig. 1). A highly conserved glycine residue lies 16-residues N-terminal to the second motif, on the edge of the predicted transmembrane segment, and might be part of a three-dimensional flexible hinge that gives mobility to the aromatic residues.

In addition to the three conserved motifs, the RhtB proteins show additional similarity: all are hydrophobic, and their transmembrane segments (predicted by the PHDhtm program[18]) exhibit similar patterns. We propose that they belong to a new, widespread class of functionally important transporters that allow excretion of metabolites from different prokaryotes and archaea.

Acknowledgements

We thank T. P. Turova and R. L. Tatusov for the SEQEDIT manual alignment software, and D. G. Naumoff and M. S. Gelfand for reading the manuscript and for discussions.

References

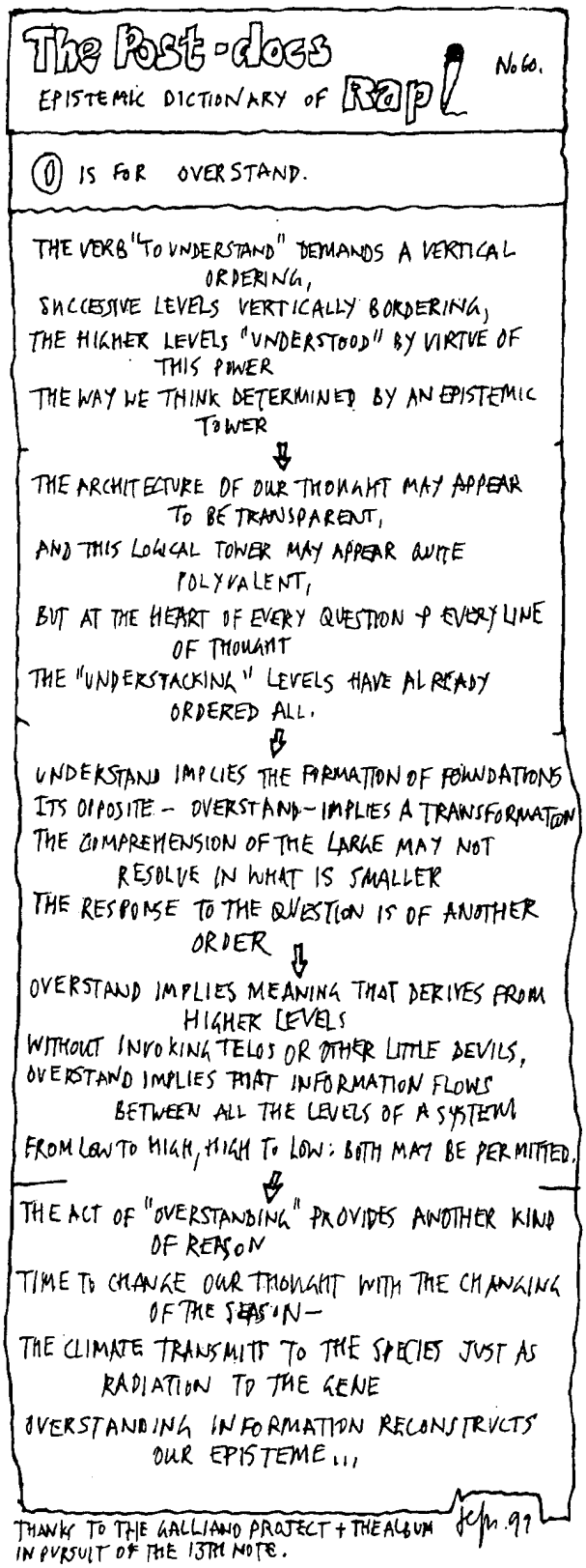
1 Griffith, J. K. *et al.* (1992) *Curr. Biol.* 4, 684–695
2 Saier, M. H., Jr (1994) *Microbiol. Rev.* 58, 71–93
3 Paulsen, I. T., Brown, M. H. and Skurray, R. A. (1996) *Microbiol. Rev.* 60, 575–608
4 Krämer, R. (1994) *FEMS Microbiol. Rev.* 13, 75–93
5 Fuqua, C., Winans, S. C. and Greenberg, E. P. (1996) *Annu. Rev. Microbiol.* 50, 727–751
6 Kotre, A. M., Sullivan, S. J. and Savageau, M. A. (1973) *J. Bacteriol.* 116, 663–672
7 Huisman, G. W. and Kolter, R. (1994) *Science* 265, 537–539
8 Loewen, P. and Hengge-Aronis, R. (1994) *Annu. Rev. Microbiol.* 48, 53–80
9 Furukawa, H. *et al.* (1993) *J. Bacteriol.* 175, 3723–3729
10 Nikaido, H. (1996) *J. Bacteriol.* 178, 5853–5859
11 Carlin, A. *et al.* (1995) *J. Bacteriol.* 177, 981–986
12 Zakataeva, N. P., Aleshin, V. V. and Livshits, V. A. (1997) *FASEB J.* 11, A935
13 Daniels, D. L. *et al.* (1992) *Science* 257, 771–778
14 Altschul, S. F. *et al.* (1997) *Nucleic Acids Res.* 25, 3389–3402
15 Vrijic, M., Sahm, H. and Eggeling, L. (1996) *Mol. Microbiol.* 22, 815–826
16 Felsenstein, J. (1985) *Evolution* 39, 783–791
17 Schuler, G. D., Altschul, S. F. and Lipman, D. J. (1991) *Protein Struct. Funct. Genet.* 9, 180–190
18 Rost, B., Casadio, R., Fariselli, P. and Sander, C. (1995) *Protein Sci.* 4, 521–533

VLADIMIR V. ALESHIN

State Research Institute of Genetics and Selection of Industrial Microorganisms, Moscow 113545, Russia; and Laboratory of Microorganism Biotechnology, Research Institute of Physiology, Biochemistry and Nutrition of Livestock, Borovsk, Kaluga region, 249010, Russia. Email:petr@bioevol.genebee.msu.su

NATALIA P. ZAKATAEVA AND VITALIY A. LIVSHITS

State Research Institute of Genetics and Selection of Industrial Microorganisms, Moscow 113545, Russia.



Pete Jeffs is a freelancer working in Paris, France.